4.0 METEOROLOGICAL DATA

This section describes the meteorological data and processing techniques used by EPA. Several aspects of EPA's CALMET meteorological data processing warrant further consideration and possible refinements. Below is a discussion of CALMET options that may result in an unrepresentative meteorological dataset for the modeling study, and could result in inaccurate model results.

Use of BIAS Factors in CALMET Processing

In the IWAQM scheme, both surface and upper air data are given equal weight when determining the wind fields at all heights. In the revised scheme adopted by EPA and NDDH, the surface stations have relatively more influence at the surface, and less influence aloft, while the upper air stations have the opposite effect.

The intent of the BIAS factor adjustments used by both EPA and NDDH is reasonable. It is appropriate for the surface monitors to have the greatest influence on the surface winds, while the upper air data would have the greatest influence on winds aloft. However, there appears to be no basis for the actual BIAS values selected in the EPA study. Also, surface topographic features would be expected to have some influence aloft but the dependence has been set to 1.0 for the top three layers, essentially indicating that there is no influence. It is uncertain whether the BIAS factor adjustments selected by EPA were based upon sound theoretical considerations, or were selected to arrive at satisfactory limited evaluation results.

5.0 RECEPTORS

This section describes the receptors used by EPA for their January 2002 modeling analysis of the Class I areas in western North Dakota. EPA's level of technical and regulatory appropriateness for using the chosen receptors is addressed, and some alternative methodologies are presented.

5.1 RECEPTOR LOCATIONS

EPA used groups of receptors to cover each of the Class I areas in the modeling analysis. There were 22 receptors placed in the TRNP South Unit, 16 receptors in the TRNP North Unit, one receptor in the TRNP Elkhorn Unit, five receptors in the Lostwood Wilderness Area, one receptor in the Medicine Lake Wilderness Area, and four receptors in the Fort Peck Reservation. Receptor spacing for Medicine Lake and Fort Peck Class I areas was less dense than receptor spacing in the other Class I areas. Medicine Lake and Fort Peck are located further away from the largest contributing sources, and not all local minor sources could be identified. According to EPA, there are fewer emissions in close vicinity to the Medicine Lake and Fort Peck areas, so receptor spacing was increased because fewer impacts were expected.

5.2 SPATIAL AVERAGING

One alternative method of estimating expected concentration impacts from distant sources is to average the concentrations over an area expected to represent the broad plume impact in the Class I area. EPA has adopted the concept of spatial averaging for other long-range transport pollutant issues, such as PM-2.5. For this PSD Class I study, there are several reasons why spatial averaging is worth serious consideration.

IWAQM (1998) has noted that the CALPUFF-predicted plumes demonstrate uncertainties in their trajectories as a result of the large distances involved, and are generally expected to be too narrow at these distances. From studies referenced in the IWAQM Phase 2 report, peak predicted concentrations and their locations are not reliable at distances of 200 kilometers, and the observed concentration field is likely to be more uniform than the predicted concentration field. Due to these model uncertainties, a spatial-averaged prediction is more likely to be in good agreement with a representative observation than the peak prediction with the peak observation. If the ratio of the peak predicted CALPUFF concentration to the spatially averaged CALPUFF concentration is less than the expected CALPUFF overprediction factor (when using peak point concentration), as it is in this case, then the use of spatial averaging should provide a more realistic estimate of the SO_2 concentrations within the PSD Class I areas.

6.0 CALPUFF MODELING APPLICATION

This section presents a review of EPA's modeling application of CALPUFF. In this section, differences between EPA's and NDDH's 1999 modeling application are discussed. In addition, differences between EPA's modeling application and IWAQM guidance are summarized. A discussion of the NDDH's 2000 model performance evaluation on which EPA relied is also presented, as well as EPA's assessment of the model performance evaluation. In addition, the correct methods for including minor sources in the modeling analysis is addressed. A presentation of conclusions follows these assessments.

6.1 LIMITATIONS OF MODEL PERFORMANCE EVALUATION

The NDDH evaluation study conducted for the year 2000 can be further improved as noted below:

- Numerous oil and gas production sources beyond 50 kilometers from the monitors were not modeled. Their omission is accounted for in the estimate of regional background, discussed below.

- The monitors are both of the pulsed-fluorescent type, with a threshold detection level of 2 parts per billion (ppb). Reported zero observed values were adjusted to half the threshold value (1 ppb) in the model evaluation study. However, zero predicted values were not similarly adjusted, leading to likely underestimation of the predicted values without additional considerations.

- The EPA's Guideline on Air Quality Modeling (Appendix W of 40 CFR Part 51) states in Section 9.2 that the total predicted value should include a regional background value, in addition to the impact from modeled sources, to account for natural background and unmodeled sources. In the evaluation study, NDDH did not include the regional background component. In order to correct this omission, a review was completed of the evaluation procedures and the TRNP South Unit monitoring data for days with winds from a southerly direction (for which there are no upwind major SO_2 sources).

- The monitored values for days with southerly winds support a regional background value of at least 1 ppb. For the critical easterly wind cases with more population centers, more oil and gas sources, and Interstate 94 highway sources, a regional background of 1.5 ppb (about 4 micrograms per cubic meter [$\mu g/m^3$], and still below the instrument detection threshold) is reasonable. This value is still very low and is much lower than values typically used as regional background estimates in other rural states (e.g., the Alabama Department of Environmental Management (2000) uses a background concentration of 10 $\mu g/m^3$).

When a regional background of only 4 $\mu g/m^3$ is added to the model predictions, the plots of the model evaluation results change significantly from the figures shown in NDDH's model evaluation report, as seen in Figures 6-1 and 6-2 for the Dunn Center monitor and Figures 6-3 and 6-4 for the TRNP South Unit monitor. The EPA concern about model underpredictions is no longer valid (EPA's concerns were misplaced anyway because the underprediction magnitude was, at most, only about 1 $\mu g/m^3$, well below

FIGURE 6-1

MODIFIED PLOT OF CALPUFF MODEL PERFORMANCE FOR YEAR 2000 (3-HOUR AVERAGES AT DUNN CENTER)

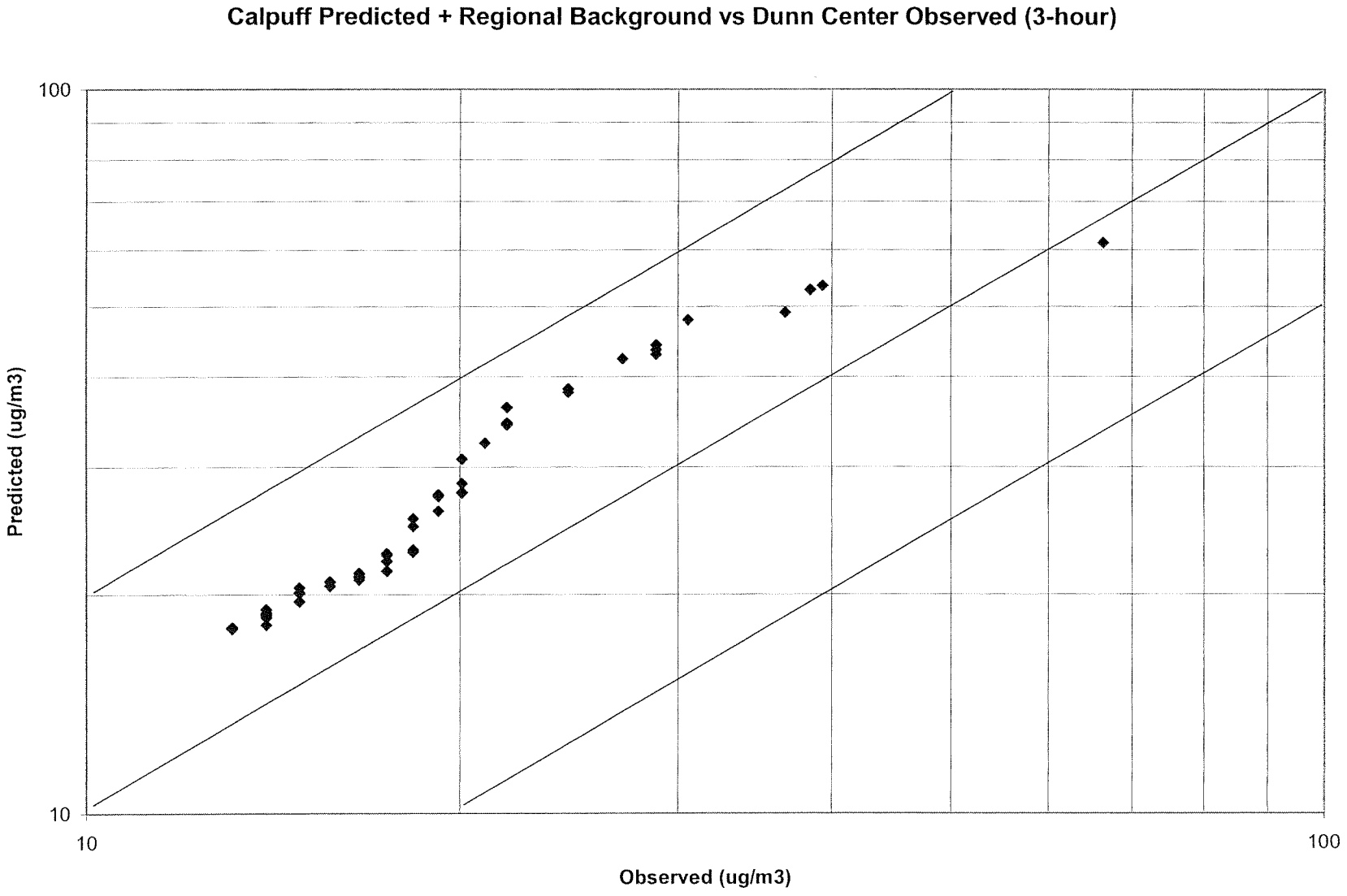


FIGURE 6-2

MODIFIED PLOT OF CALPUFF MODEL PERFORMANCE FOR YEAR 2000 (24-HOUR AVERAGES AT DUNN CENTER)

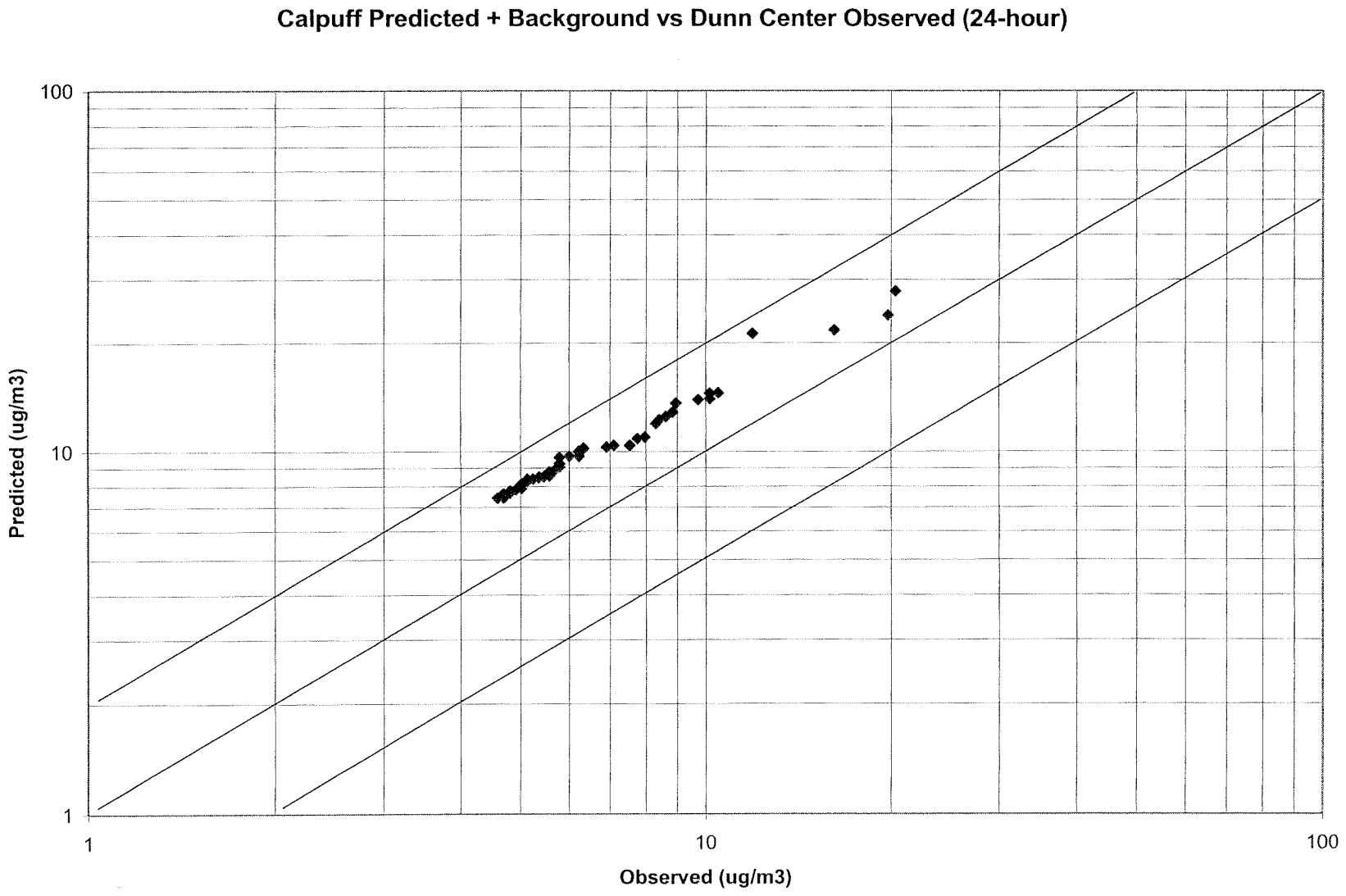


FIGURE 6-3

MODIFIED PLOT OF CALPUFF MODEL PERFORMANCE FOR YEAR 2000 (3-HOUR AVERAGES AT TRNP SOUTH UNIT)

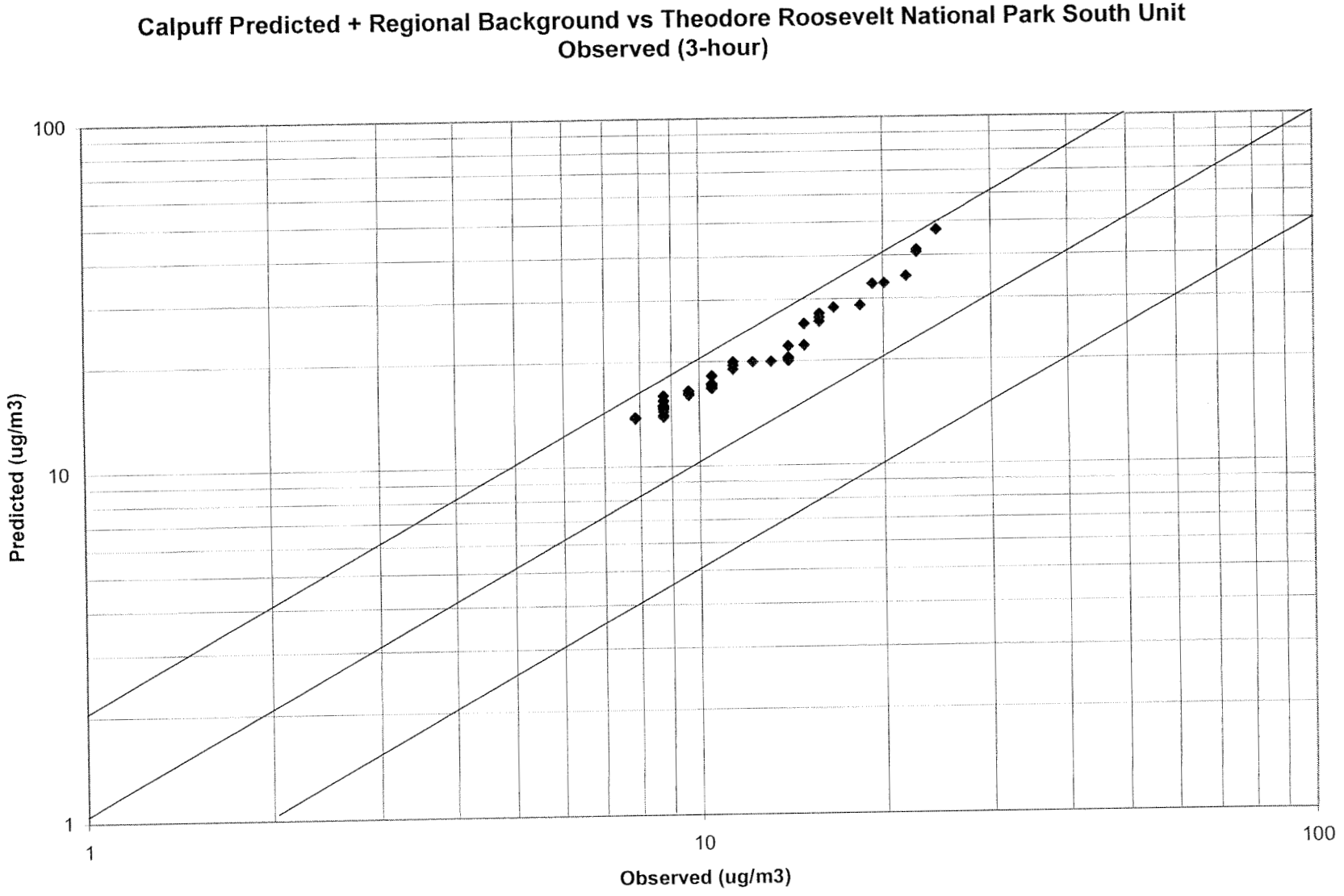
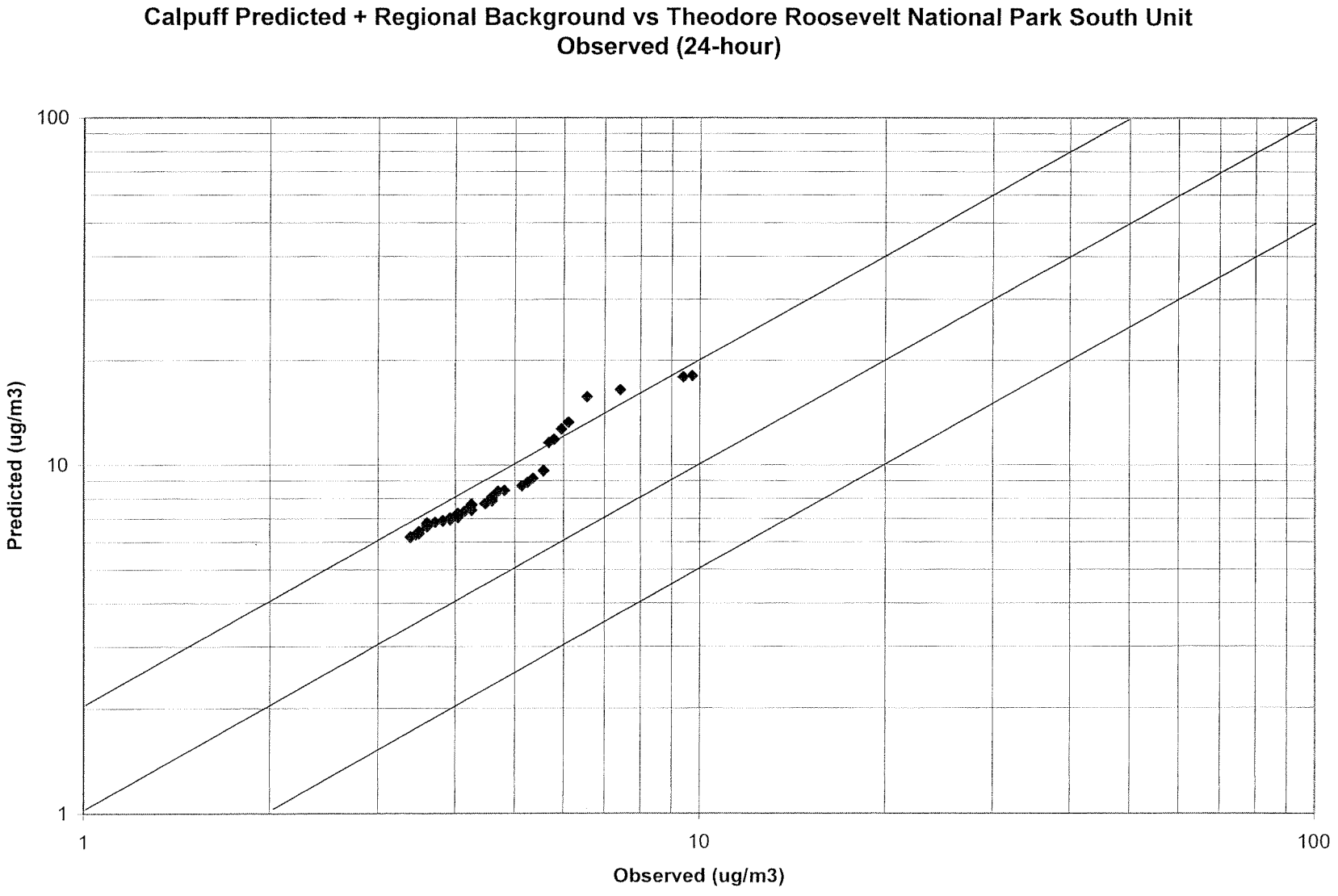


FIGURE 6-4

MODIFIED PLOT OF CALPUFF MODEL PERFORMANCE FOR YEAR 2000 (24-HOUR AVERAGES AT TRNP SOUTH UNIT)



the instrument threshold). The new results indicate that for Dunn Center (roughly 100 kilometers from many of the major sources), CALPUFF overpredicts on average by roughly 50 percent for the top several concentrations. For TRNP South Unit (roughly 150 to 200 kilometers from many of the major sources), the CALPUFF overprediction tendency for the peak 3-hour concentrations is nearly 2.0, and it slightly exceeds 2.0 for the highest 24-hour averages.

6.2 MODELED VERSUS MEASURED DATA, PARING AND UNPARING IN TIME

EPA has a long-standing and documented policy concerning the conduct of model evaluation using data that are unpaired in time. The argument centers on the predicted narrowness of plumes and the inability to identify plume trajectories well enough to determine when a plume will impact a set of receptors. The failure to accurately predict concentrations when paired in time, particularly when multiple sources of a pollutant are present, is a problem that needs to be considered in the interpretation and application of model results.

In the case of EPA's CALPUFF analysis, when concentrations are paired in time, the performance of the model is poor. For example, compare EPA's model results (using average emissions representative of the 2000-2001 period) with measured data for the first 6 months of data at TRNP south unit for 1990 (the monitor was discontinued after the first half of 1990). The paired in time comparison is demonstrated in Figure 6-5.

Perfect performance would line up on the diagonal running from the lower left corner to the upper right corner. As the figure shows, actual performance of the model was quite poor. In fact, for the highest measured hour, the model predicted zero, and for most of the high-predicted hours, the measured value was zero. In fact, for all measured values higher than 20 $\mu g/m^3$, the model prediction was less than 0.2 $\mu g/m^3$. Similarly, of the top 40 model-predicted values, only 2 had non-zero measured values. The model performance, when viewed paired in time, is very poor.

In contrast, good dispersion models do have skill in predicting concentrations paired in space, but unpaired in time. This is strikingly evident in Figure 6-6, which takes the data shown in Figure 6-5 and simply plots the pairs of ranked concentrations unpaired in time.

The limitations of the ability for a dispersion model to predict well for specific events at specific locations, but to match the ranked concentrations unpaired in time have important policy implications. In most PSD applications, the increment concentration expansion due to sources shutting down or reducing

FIGURE 6-5

THEODORE ROOSEVELT NATIONAL PARK SOUTH UNIT - 1990 (JANUARY - JUNE)
1-HOUR SO_2 CONCENTRATIONS- PAIRED IN TIME

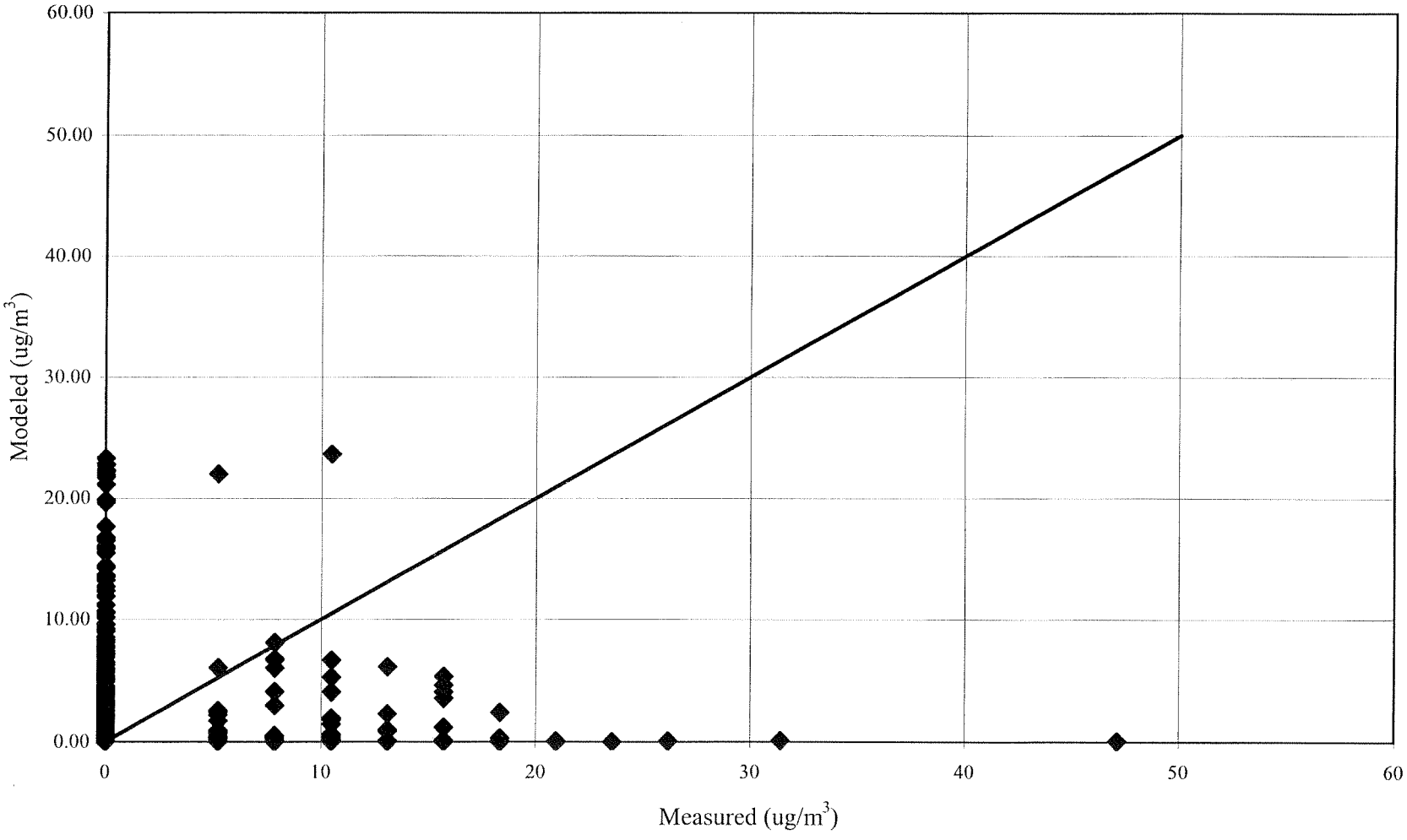
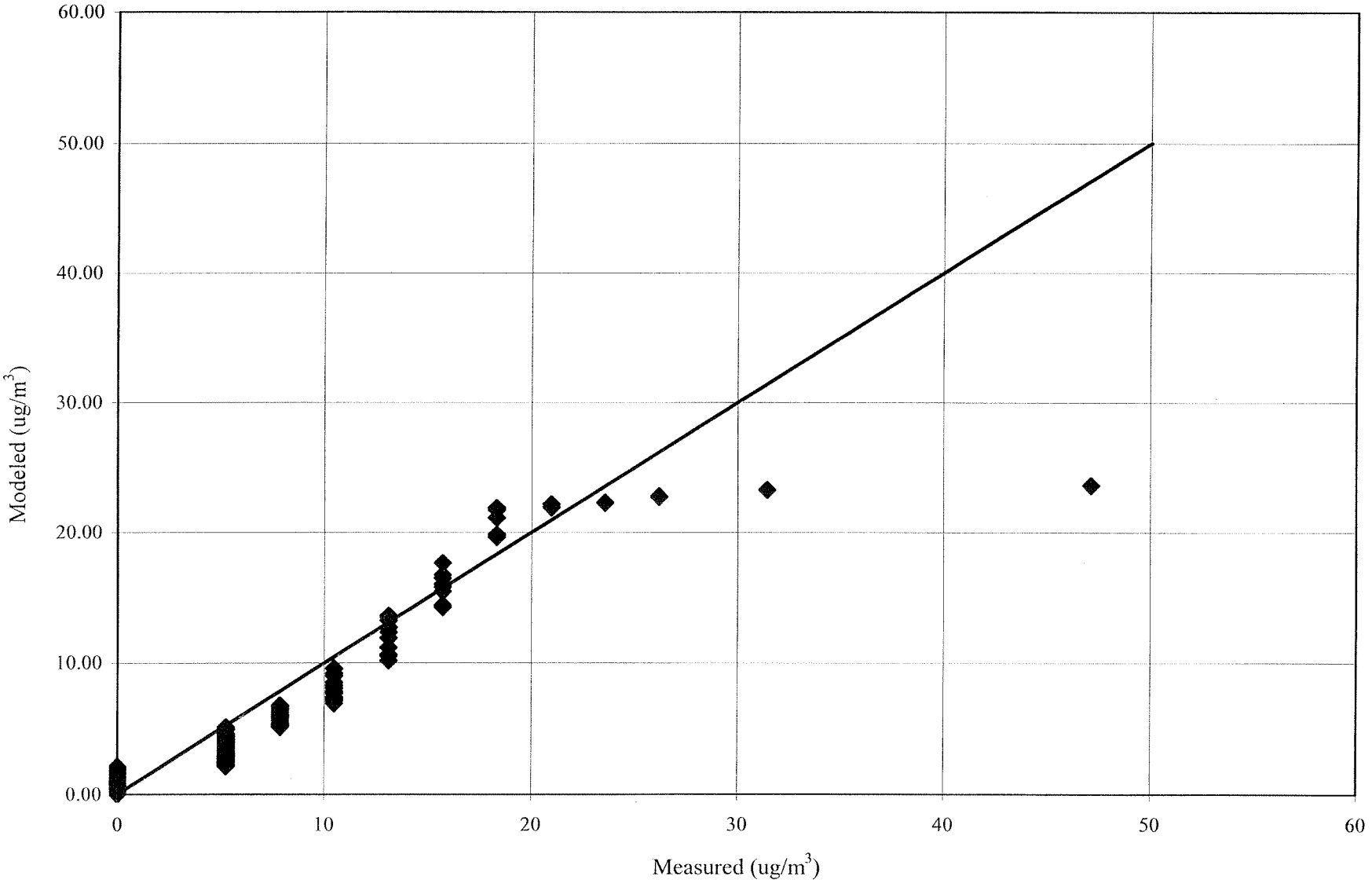


FIGURE 6-6

THEODORE ROOSEVELT NATIONAL PARK SOUTH UNIT - 1990 (JANUARY - JUNE)
1-HOUR SO_2 CONCENTRATIONS - UNPAIRED IN TIME



emissions fails to cancel the increment consumption from new sources (however clean) because the two plumes do not coincide in time and space, according to the dispersion model.

EPA's procedure for calculating increment in the current project has this problem. Because EPA determines increment by subtracting the baseline concentrations from the current concentrations on an hour-by-hour basis, they virtually guarantee that North Dakota will fail the increment test. There will inevitably be hours when the plumes from the baseline sources impact different receptors than the current sources. So even if the overall air quality in the Class I area stays the same or even improves, EPA's method will lead to a conclusion that increment is exceeded, because changes at a particular receptor for a particular 3-hour or 24-hour period will exceed the criteria. This technique and conclusion are only valid if the model has demonstrated skill at predicting concentrations at a particular location and time, but it has not.

It is evident from Figure 6-5 that CALPUFF (similar to all other dispersion models) is incapable of accurately determining where a plume will impact during a given event. The task of accurately determining the locations of multiple plumes (some increment consuming and some increment expanding) would be even more difficult for a dispersion model to perform. As Figure 6-6 shows, the model only demonstrates any measure of skill when measured and modeled concentrations are unpaired in time. In effect the model states that it cannot predict when a concentration will occur, but rather can estimate the statistical distributions of concentrations at a particular location. To accommodate the limitations of dispersion models, the subtraction of baseline concentrations from current concentrations must not be paired in time. This goal can be achieved by first modeling the baseline emissions for a given model receptor and then determining the controlling concentration from a regulatory point of view (the second highest short-term average). The same procedure would be separately done for the current/future emissions, depending upon the application. The difference of the unpaired (time-independent) second highest concentrations from the two runs would be used to determine the increment that is consumed.

6.3 MINOR SOURCES EVALUATION

EPA failed to include minor sources in the analysis. Significant benefit to air quality may have resulted from increased controls on many of the minor emission sources in the area since the baseline date that EPA has failed to include in the analysis. The conclusions on increment consumption may be wrong.

An ISCST3 modeling analysis was performed for the minor sources near TRNP South Unit for 1990. Included in the modeling were only those sources identified by NDDH and included in the year 2000

model evaluation runs. Model runs were made for two cases: a baseline case, and a "current" case, termed an Y2K case by NDDH. The ISCST3 modeling showed significant increment expansion from the reduction in emissions. At some receptors in the TRNP South unit, peak 24-hour SO_2 concentrations were reduced by over 8 $\mu g/m^3$. EPA did not include minor sources in their modeling analysis, and therefore did not address the increment expansion that may have occurred as a result of these emission reductions.

6.4 CONCLUSIONS AND RECOMMENDATIONS

EPA's modeling application of the CALPUFF model raises issues with (1) the comparison of EPA's and NDDH's model application, (2) the comparison of EPA's model application with IWAQM guidance, (3) EPA's analysis of NDDH's 2000 model performance evaluation, and (4) EPA's failure to include minor source in their analysis. The potential problems identified with these issues could demonstrate inconsistencies with EPA's modeled results.

EPA differed with NDDH on the values used for several modeled parameters such as mixing height, background ammonia concentrations, and interval distance between puffs. These differences, while at times small, may have an impact on the results.

At other times, EPA and NDDH used identical values for modeled parameters but both modeled applications differed from IWAQM guidance. These variations from IWAQM guidance include differences in model settings such as background ozone concentrations and plume splitting. These distinctive model settings were selected because they reportedly improved model performance when compared with measured data. However, it has not been demonstrated that these settings are representative of actual conditions.

In terms of model performance evaluation, the procedures used by the NDDH in their 2000 study corrected deficiencies in the earlier study reported in 1999, which did not use hourly emission rates from the major sources. In addition, the emissions data in 2000 used a more accurate flow measurement technique that avoided overestimates of stack emissions associated with methods used prior to 2000. Both model performance evaluations omitted regional background, and in doing so, failed to accurately depict the clear overprediction tendency of CALPUFF, demonstrated to be on the order of twice the measured levels. If the performance evaluation is modified to correctly apply this adjustment, the overprediction tendency is clearly demonstrated.

These modified model evaluation results are consistent with the IWAQM Phase 2 report findings (noted in Section 2 of this report). They demonstrate a likely CALPUFF model overprediction at the distances being considered for this modeling application. With corroboration from this limited evaluation study, the EPA modeling results are therefore likely to be subject to the same overprediction problem, and the findings from the EPA study must be viewed with these overprediction tendencies in mind. Either the EPA modeling procedures need to be corrected to eliminate the overprediction tendency, or the results need to be adjusted to account for the overprediction.

In addition, changes should be considered for EPA's model analysis. A more comprehensive evaluation of oil and gas emission sources should be included. For those sources within 50 kilometers of the Class I areas, EPA model guidance should be followed. The CALPUFF model is only proposed for long-range transport applications. For applications where the sources are within 50 kilometers of the receptors of interest, the recommended model remains ISCST3, although AERMOD is presently proposed to replace ISCST3 in many of the applications.

7.0 MODEL RESULTS

The model results listed in Tables 4-1 through 4-6 of EPA's report represent the highest and "high, second-highest" modeled concentrations at each Class I area for each modeled year and averaging period. The term "maximum number of exceedances" listed in the tables is confusing and should be more clearly defined. The value apparently represents the highest number of concentrations that exceed the standard at any one receptor. However, the "maximum number of exceedances" term as used in Tables 4-1 through 4-6 is not the same as the "number of violations" term presented in Table 5-1 of EPA's report. Since one exceedance of the PSD standard is allowed per year at each receptor, the number of violations should be one less than the number of exceedances for each receptor.

The results presented in Tables 4-1 through 4-6 of EPA's report also do not correspond with post-processing output files contained on the modeling compact disk (CD) that EPA distributed with the report. For example, the highest 3-hour modeled concentration for 1990 at TRNP South Unit listed in Table 4-1 of EPA's report is 36.4 $\mu g/m^3$. But the highest 3-hour concentration listed in the CALEXCEED output file (puftot90.1st) provided on CD is 40.9 $\mu g/m^3$. Similarly the highest 3-hour modeled concentration for 1990 at TRNP-North Unit in Table 4-1 of EPA's report is 29.4 $\mu g/m^3$, but is 28.6 $\mu g/m^3$ in puftot90.1st. The post-processing files may originate from preliminary model runs conducted by EPA.

8.0 CONCLUSIONS

We have reviewed EPA's report entitled Dispersion Modeling Analysis of PSD Class I Increment Consumption in North Dakota and Eastern Montana dated January 2002. We conclude that the EPA modeling method should not be used at this time for the following reasons:

- CALPUFF has not been designated as a regulatory-approved model. The proposed model used by EPA may change prior to final promulgation.
- Several problems have been identified with the emissions inventory. These include:
  - Failure to include increment expanding minor sources
  - Underestimation of certain increment expanding sources
  - Failure to exclude variance emissions
  - Failure to include increment expanding emissions from the Mandan Refinery
  - Inconsistent approach for calculating current and baseline emissions from major sources
- We have identified questions about the processing of the meteorological data. Most significantly, we are concerned about the lack of meteorological data in the region between the major SO_2 sources and the PSD Class I areas. Further the use of a coarse grid resolution, use of a modified wind extrapolation method, and, in some cases, use of arbitrary input parameters all bring question to the validity of the meteorological dataset.
- There is evidence that CALPUFF is overpredicting concentrations by about a factor of 2.
- The presentation of the model results in EPA's report was unclear and therefore confusing.
- Test model runs indicated that CALPUFF results vary significantly with changing input parameters. Therefore, additional discussion and testing of these parameters is required before conclusions can be drawn based on model runs.

As a result, primarily, of (1) lack of information and clear errors in the input parameters (such as not using emission representative of the normal operations of the sources involved, nor determining whether they were in normal operation on the minor source baseline date), (2) omission of increment expanding sources, including at least one major stationary source and all oil and gas sources, and (3) failure to include regional background in performance evaluation, and thus failure to consider the likely overprediction of the model by a factor of two, EPA's preliminary modeling is not adequate for regulatory use. It is clearly inadequate as the basis for a SIP call.

9.0 BIBLIOGRAPHY

Alabama Department of Environmental Management. 2000. PSD Air Quality Analysis Modeling Guidelines. Available from ADEM Air Quality Branch, Montgomery, Alabama.

Anthes, R.A., E.Y. Hsie, and Y.H. Kuo. 1987. Description of the Penn State/NCAR mesoscale.model version 4 (MM4). NCAR Tech Note, NCAR/TN-282+STR, Available from the National Center for Atmospheric Research, PO Box 3000, Boulder, CO 80307.

Interagency Workgroup on Air Quality Modeling. 1998. Interagency Workgroup On Air Quality Modeling (IWAQM) Phase 2 Summary Report and Recommendations for Modeling Long Range Transport Impacts. EPA-454/R-98-019.

Moran, M.D. and R.A. Pielke. 1994. Delayed Shear Enhancement in Mesoscale Atmospheric Dispersion. Paper 3.1 in proceedings of the Eighth Joint Conference on Applications of Air Pollution Meteorology with A&WA. American Meteorological Society, Boston, MA.

Morris, R.E., R.C. Kessler, S.G. Douglas, K.R. Styles, and G.E. Moore. 1988. Rocky Mountain Acid Deposition Model Assessment: Acid Rain Mountain Mesoscale Model (ARM3). EPA-600/3-88-042. U.S. Environmental Protection Agency, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC.

North Dakota Department of Health (NDDH). 1999. CALPUFF Class I Area Analysis for Milton R. Young Generating Station. Bismarck, North Dakota. May.

Scire, J.S., D.G. Strimaitis, and R.J. Yamartino. 2000. A User's Guide for the CALPUFF Dispersion Model (Version 5.4). http://www.epa.gov/scram001 (under 7th Modeling Conference link to Earth Tech web site). (accessed December 2000).

U.S. Environmental Protection Agency (EPA). 1993. Interagency Workgroup on Air Quality Modeling (IWAQM) Phase 1 Report: Interim Recommendation for Modeling Long Range Transport and Impacts on Regional Visibility. EPA-454/R-93-015. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC.

EPA. 1994. Revised User's Guide to the MESOPUFF II (V5.1) EPA-454/B-94-025. U.S. Environmental Protection Agency, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC.